Linear Algebra and its Applications 433 (2010) 1241–1253

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Third-order tensors as linear operators on a space of matrices

Karen Braman

Department of Mathematics and Computer Science, South Dakota School of Mines and Technology, 501 E. St Joseph St., Rapid City, SD 57701, USA

ARTICLE INFO

Article history:
Received 4 June 2009
Accepted 15 April 2010
Available online 6 July 2010

Submitted by V. Mehrmann

AMS classification:
15A69

Keywords:
Multilinear algebra
Tensor decomposition
Diagonalization

ABSTRACT

A recently proposed tensor–tensor multiplication (M.E. Kilmer, C.D. Martin, L. Perrone, *A Third-Order Generalization of the Matrix SVD as a Product of Third-Order Tensors*, Tech. Rep. TR-2008-4, Tufts University, October 2008) opens up new avenues to understanding the action of $n \times n \times n$ tensors on a space of $n \times n$ matrices. In particular it emphasizes the need to understand the space of objects upon which tensors act. This paper defines a free module and shows that every linear transformation on that module can be represented by tensor multiplication. In addition, it presents a generalization of ideas of eigenvalue and eigenvector to the space of $n \times n \times n$ tensors.

1. Introduction

> What exactly are the objects that are operated on by tensors? What are the singular values or eigenvalues? Could they be vectors rather than scalars? [1]

Those questions are among a list of "open questions related to extending linear algebra theory" generated by the participants at the American Institute of Mathematics "Workshop on Tensor Decompositions" held in 2004 [2,3]. The term *tensor*, as used above and in the context of this paper, refers to a multi-dimensional array of numbers, sometimes called an *n-way* or *n-mode* array. If, for example, $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ then we say $\mathcal{A}$ is a third-order tensor where *order* is the number of ways or modes of the tensor. Thus, matrices and vectors are second-order and first-order tensors, respectively. Restricting our attention to third-order tensors where $n_1 = n_2 = n_3$, we will present answers to the three of the four questions posed above. (We will not investigate singular values in this paper.)

E-mail address: Karen.Braman@sdsmt.edu

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.05.025

Table 1
Summary of notation for third-order tensors.

$\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$	
$\mathcal{A}_{ijk}$	The (i,j,k) element of $\mathcal{A}$
$\mathcal{A}_{:jk}$	A row of $\mathcal{A}$
$\mathcal{A}_{i:k}$	A column of $\mathcal{A}$
$\mathcal{A}_{ij:}$	A tube of $\mathcal{A}$
$\mathcal{A}_{i::}$	A horizontal slice of $\mathcal{A}$
$\mathcal{A}_{:j:}$	A lateral slice of $\mathcal{A}$
$\mathcal{A}_{::k}$	A frontal slice of $\mathcal{A}$

This work has been inspired by, and builds upon, a recent paper from Kilmer, Martin and Perrone where they define a new tensor–tensor multiplication and present a procedure for producing tensor decompositions relative to that operation [4]. Their work focuses on producing a tensor SVD, but, as they point out, the process can be modified to compute a tensor QR decomposition or even a *tensor diagonalization*. Given a tensor $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$, under certain conditions, we can compute tensors $\mathcal{X}$ and $\mathcal{L} \in \mathbb{R}^{n\times n\times n}$ such that

$$\mathcal{A} * \mathcal{X} = \mathcal{X} * \mathcal{L} \tag{1}$$

where $*$ denotes the new tensor–tensor multiplication and $\mathcal{L}$ is *f-diagonal*, that is, each frontal-face slice is a diagonal matrix. The existence of this diagonalization underscores the importance of the questions cited above.

In Section 2 we establish the notation and definitions needed throughout the paper. In particular, we reiterate the tensor–tensor multiplication defined in [4] and specialize from it a definition of tensor–matrix multiplication. In the first of several parallels to matrix theory, we state and prove a lemma relating tensor–tensor multiplication to tensor–matrix multiplication. From this result we see that $n \times n \times n$ tensors must act upon a space of $n \times n$ matrices. Section 3 begins with an explanation of why the usual vector space over $\mathbb{R}$ cannot be the appropriate space in which to view the action of tensor–matrix multiplication. The remainder of the section is spent constructing a space where the vectors are $n \times n$ matrices and the scalars are vectors of length n with a carefully chosen multiplication between such objects. In Section 4 we explore the action of a given $n \times n \times n$ tensor upon elements of that space, and, in particular, we establish that every linear operator on the space can be represented by tensor multiplication. Section 5 examines the implications of Eq. (1) in more detail and generalizes the notions of eigenvalue and eigenvector to third-order tensors. Finally, in Section 6 we consider further avenues of research.

2. Notation and definitions

2.1. Notation

Throughout this paper tensors are denoted with Euler script letters (e.g., $\mathcal{A}$), while capital letters represent matrices (e.g., I denotes the identity matrix), boldface lowercase letters represent vectors, and lowercase letters refer to scalars. We use Matlab style notation to refer to columns, rows, tubes, and various two-dimensional slices of third-order tensor as outlined in Table 1.

Several definitions and theorems in this paper use circulant or block-circulant matrices defined by column-wise. Specifically, if

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}^T$$

then

$$circ\,(\mathbf{v}) = \begin{bmatrix} v_1 & v_n & \cdots & v_2 \\ v_2 & v_1 & \cdots & v_3 \\ \vdots & \vdots & & \vdots \\ v_n & v_{n-1} & \cdots & v_1 \end{bmatrix}$$

As it is often necessary to refer to individual rows or columns in a circulant matrix, we define the following notation:

Indices for the jth *row:*

$$row(j,k) \equiv ((j-k) \bmod n) + 1, \quad \text{for } k = 1, 2, \ldots, n$$

Indices for the jth *column:*

$$col(j,k) \equiv ((n-j+k) \bmod n) + 1, \quad \text{for } k = 1, 2, \ldots, n$$

2.2. Definitions

Fundamental to the results of this paper is a tensor–tensor multiplication recently introduced by Kilmer et al. [4]. For clarity, some of their definitions and examples are reproduced here, specifically, Items 2.1–2.4.

Definition 2.1. Let $\mathcal{A} \in \mathbb{R}^{p \times s \times r}$ and let A_i denote $\mathcal{A}_{::i}$, that is, the ith frontal slice of $\mathcal{A}$.

Define

$$\texttt{unfold}(\mathcal{A}) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{bmatrix} \in \mathbb{R}^{pr \times s}.$$

Similarly, the operation needed to reconstruct a tensor from a $pr \times s$ matrix is denoted

$$\texttt{fold}\left(\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{bmatrix}\right) = \mathcal{A} \in \mathbb{R}^{p \times r \times s}.$$

See Fig. 1.

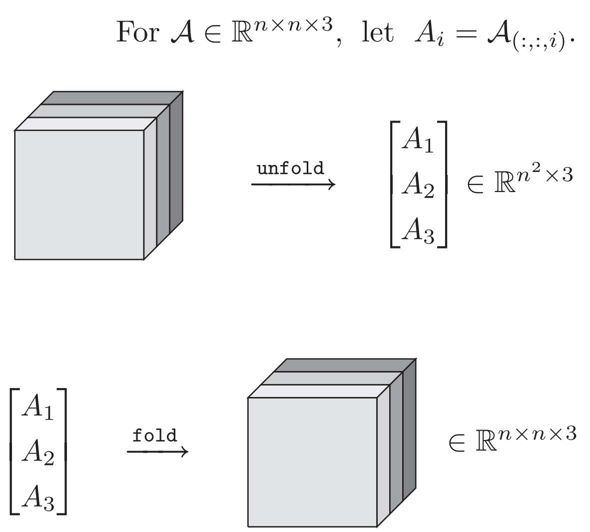

Fig. 1. Example of the `unfold` and `fold` operators.

Definition 2.2. Let $\mathcal{A} \in \mathbb{R}^{p\times s\times r}$ and $\mathcal{B} \in \mathbb{R}^{s\times q\times r}$. Then the product $\mathcal{A} * \mathcal{B}$ is the $p \times q \times r$ tensor

$$\mathcal{A} * \mathcal{B} \equiv \texttt{fold}(\mathit{circ}\ (\texttt{unfold}(\mathcal{A})) \cdot \texttt{unfold}(\mathcal{B}))$$

where $\cdot$ denotes the usual matrix–matrix multiplication. Note that $\texttt{unfold}(\mathcal{A})$ is a column of $p \times s$ matrices and so $\mathit{circ}\ (\texttt{unfold}(\mathcal{A}))$ results in a block circulant matrix of size $pr \times sr$.

In the following example, we again let A_i and B_i denote $\mathcal{A}_{::i}$ and $\mathcal{B}_{::i}$, that is, the ith frontal slices of $\mathcal{A}$ and $\mathcal{B}$, respectively.

Example 2.3. Suppose $\mathcal{A} \in \mathbb{R}^{n\times n\times 3}$ and $\mathcal{B} \in \mathbb{R}^{n\times n\times 3}$. Then

$$\mathcal{A} * \mathcal{B} = \texttt{fold}\left(\begin{bmatrix} A_1 & A_3 & A_2 \\ A_2 & A_1 & A_3 \\ A_3 & A_2 & A_1 \end{bmatrix}\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}\right).$$

Definition 2.4. The $n \times n \times p$ identity tensor $\mathcal{I}_{nnp}$ is the tensor whose front face is the $n \times n$ identity matrix I_n and whose other faces are all zeros. That is,

$$\mathcal{I}_{nnp} = \texttt{fold}\left(\begin{bmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right).$$

Just as matrix–matrix multiplication $A\ B$ can be viewed as the matrix A acting upon the columns of B, we propose to show a tensor–tensor multiplication counterpart. Toward that end, we define a tensor–matrix multiplication. First, note that the set of $n \times n$ matrices can be identified with the set of all $n \times 1 \times n$ tensors in an obvious way. That is, we consider a matrix $B \in \mathbb{R}^{n\times n}$ as a tensor oriented as a single lateral slice:

$$\mathcal{B}_{i1j} = B(i,j) \quad \text{for } i,j = 1,\ldots,n$$

so

$$\mathcal{B}_{:1:} = B.$$

Tensor–matrix multiplication then follows as a special case of the earlier definition.

Definition 2.5. Let $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$ and let $B \in \mathbb{R}^{n\times n}$. Then the product $\mathcal{A} * B$ is the $n \times 1 \times n$ tensor or, equivalently, the $n \times n$ matrix:

$$\mathcal{A} * B \equiv \texttt{fold}\left(\mathit{circ}\ (\texttt{unfold}(\mathcal{A})) \cdot \texttt{unfold}(\mathcal{B})\right).$$

To simplify notation we will shorten $\mathit{circ}\ (\texttt{unfold}(\mathcal{A}))$ to $\mathit{circ}\ (\mathcal{A})$ and, when $\mathcal{B}$ is really a matrix, $\texttt{unfold}(\mathcal{B})$ to $\mathit{vec}(B)$. Note that, by the action of the $\texttt{fold}$ operator on $\mathit{circ}\ (\mathcal{A}) \cdot \mathit{vec}(B)$, each column of the matrix $\mathcal{A} * B$ is a sum of matrix–vector multiplies of faces of $\mathcal{A}$ with columns of B. Specifically,

$$(\mathcal{A} * B)_{:i} = \sum_{k=1}^{n} A_{row(i,k)}\mathbf{b}_k,$$

where $A_{row(i,k)}$ is the frontal slice of $\mathcal{A}$ located in the kth position of the ith row of $\mathit{circ}\ (\mathcal{A})$ and $\mathbf{b}_k$ is the kth column of B.

Example 2.6. Again, suppose $\mathcal{A} \in \mathbb{R}^{3\times 3\times 3}$, but let $B \in \mathbb{R}^{3\times 3}$. When B is considered as a lateral tensor, its frontal slices are simply the columns of B. Let $\mathbf{b}_i$ denote the ith column of B. Then

$$\mathcal{A} * B = \texttt{fold}\left(\mathit{circ}\ (\mathcal{A})\, \texttt{unfold}(B)\right) = \texttt{fold}\left(\begin{bmatrix} A_1 & A_3 & A_2 \\ A_2 & A_1 & A_3 \\ A_3 & A_2 & A_1 \end{bmatrix}\begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}\right)$$

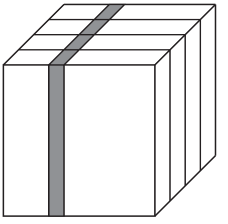

Fig. 2. $\mathcal{A}_{:i:}$ is an $n \times n$ matrix whose jth column is $\mathcal{A}_{:ij}$.

and, e.g.,

$$(\mathcal{A} * B)_{:2} = \sum_{k=1}^{3} A_{row(2,k)} \mathbf{b}_k = A_2 \mathbf{b}_1 + A_1 \mathbf{b}_2 + A_3 \mathbf{b}_3.$$

Using this definition, the following lemma illuminates the correspondence between tensor–tensor multiplication and matrix–matrix multiplication.

Lemma 2.7. *Let* $\mathcal{A}$ and $\mathcal{B} \in \mathbb{R}^{n \times n \times n}$. *Then*

$$(\mathcal{A} * \mathcal{B})_{:i:} = \mathcal{A} * (\mathcal{B})_{:i:} \quad \text{for } i = 1, \ldots, n.$$

That is,

$$\mathcal{A} * \left[\ \mathcal{B}_{:1:} \mid \mathcal{B}_{:2:} \mid \ \cdots \ \mid \mathcal{B}_{:n:}\ \right] = \left[\ \mathcal{A} * (\mathcal{B})_{:1:} \mid \mathcal{A} * (\mathcal{B})_{:2:} \mid \ \cdots \ \mid \mathcal{A} * (\mathcal{B})_{:n:}\ \right].$$

Proof. First note, for a third-order tensor $\mathcal{A}$, the ith *lateral* slice of $\mathcal{A}$ is the `fold` operator applied to the vector containing the ith columns of each frontal slice. See Fig. 2. Specifically,

$$(\mathcal{A} * \mathcal{B})_{:i:} = \texttt{fold}\left(\begin{bmatrix} (\mathcal{A} * \mathcal{B})_{:i1} \\ (\mathcal{A} * \mathcal{B})_{:i2} \\ \vdots \\ (\mathcal{A} * \mathcal{B})_{:in} \end{bmatrix} \right). \tag{2}$$

From the definition of $\mathcal{A} * \mathcal{B}$, each frontal slice $(\mathcal{A} * \mathcal{B})_{::j}$ is a sum of matrix products:

$$(\mathcal{A} * \mathcal{B})_{::j} = \sum_{k=1}^{n} \mathcal{A}_{::row(j,k)} \mathcal{B}_{::k}$$

Then the ith column of each frontal slice is the sum of the ith columns of the matrix products and so, by considering the matrix multiplication column-wise,

$$(\mathcal{A} * \mathcal{B})_{:ij} = \sum_{k=1}^{n} \left(\mathcal{A}_{::row(j,k)} \mathcal{B}_{::k} \right)_{:i} = \sum_{k=1}^{n} \mathcal{A}_{::row(j,k)} \mathcal{B}_{:ik}. \tag{3}$$

Combining Eq. (2) with Eq. (3) produces

$$\begin{aligned}
(\mathcal{A} * \mathcal{B})_{:i:} &= \texttt{fold}\left(\begin{bmatrix} \sum_{k=1}^{n} \mathcal{A}_{::row(1,k)} \mathcal{B}_{:ik} \\ \sum_{k=1}^{n} \mathcal{A}_{::row(2,k)} \mathcal{B}_{:ik} \\ \vdots \\ \sum_{k=1}^{n} \mathcal{A}_{::row(n,k)} \mathcal{B}_{:ik} \end{bmatrix} \right) \\
&= \texttt{fold}\left(\begin{bmatrix} \mathcal{A}_{::1} & \mathcal{A}_{::n} & \cdots & \mathcal{A}_{::2} \\ \mathcal{A}_{::2} & \mathcal{A}_{::1} & \cdots & \mathcal{A}_{::3} \\ \vdots & \vdots & & \vdots \\ \mathcal{A}_{::n} & \mathcal{A}_{::(n-1)} & \cdots & \mathcal{A}_{::1} \end{bmatrix} \begin{bmatrix} \mathcal{B}_{:i1} \\ \mathcal{B}_{:i2} \\ \vdots \\ \mathcal{B}_{:in} \end{bmatrix} \right) \\
&= \mathcal{A} * (\mathcal{B})_{:i:} \qquad \square
\end{aligned}$$

Analogously, *matrix–tensor* multiplication is defined by considering an $n \times n$ matrix as a tensor with a single $n \times n$ *horizontal* slice and then using the definition of $*$ given above. The lemma also has an analogous version, stated below. The proof, which is very similar, is omitted.

Lemma 2.8. *Let* $\mathcal{A}$ and $\mathcal{B} \in \mathbb{R}^{n\times n\times n}$. *Then*

$$(\mathcal{A})_{i::} * \mathcal{B} = (\mathcal{A} * \mathcal{B})_{i::} \qquad \text{for } i = 1, \ldots, n.$$

When applied to Eq. (1), Lemma 2.7 shows that, for $1 \leqslant i \leqslant n$,

$$\mathcal{A} * X_{:i:} = \mathcal{X} * L_{:i:}.$$

This highlights the need to accurately understand the action of third-order tensors upon the space of $n \times n$ matrices.

3. A free module

While it is clear from Lemma 2.7 that $n \times n \times n$ tensors act upon $n \times n$ matrices, considering the usual identification of $\mathbb{R}^{n\times n}$ with $\mathbb{R}^{n^2}$ does not provide the correct underlying space for understanding this action. This stems from the fact that, even though the new tensor multiplication is a linear operator upon $\mathbb{R}^{n^2}$, there exist linear transformations over that space which can not be represented by tensor–matrix multiplication. In addition, since the new tensor–tensor multiplication is carefully defined to avoid the loss of information inherent in *flattening* the tensors, it is desirable to maintain that perspective when considering the action of tensors upon matrices.

Thus, we require a space where the where the 'vectors' are the set of $n \times n$ matrices (not vectorized) and where the 'scalars' and 'scalar-multiplication' combine to encapsulate the behavior of tensor–matrix multiplication as a linear operator. Toward this end, we first construct a commutative ring with unity and define a free module over it.

Definition 3.1. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Define a multiplication $\odot$ on $\mathbb{R}^n$ as follows:

$$\begin{aligned}\mathbf{a} \odot \mathbf{b} &\equiv circ\,(\mathbf{a})\,\mathbf{b} \\ &\equiv \begin{bmatrix} a_1 & a_n & \cdots & a_2 \\ a_2 & a_1 & \cdots & a_3 \\ \vdots & \vdots & & \vdots \\ a_n & a_{n-1} & \cdots & a_1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.\end{aligned}$$

Theorem 3.2. $(\mathbb{R}^n, +, \odot)$ *is a commutative ring with unity, where* $+$ *denotes the usual addition of vectors.*

Proof. It is well known that $(\mathbb{R}^n, +)$ is an abelian group and it is straightforward, if tedious, to verify that $\odot$ is associative, commutative and distributes over $+$ from both the left and right. Finally, let $\mathbf{e}$ denote the first column of the identity matrix I_n. Then it is also easy to see that for any $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \odot \mathbf{e} = \mathbf{e} \odot \mathbf{a} = \mathbf{a}$. $\square$

We will use $\mathbb{K}_n$ to denote this ring.

Ultimately elements of $\mathbb{K}_n$ will play the role of 'scalars' in our new space, so it would be ideal for $\mathbb{K}_n$ to be a field. Unfortunately, that is not the case. The next example shows, for $n = 2$, $\mathbb{K}_n$ is not a domain, let alone a field.

Example 3.3. Consider $\mathbb{K}_2$ and let $\mathbf{a} = [\,a \quad a\,]^T \in \mathbb{K}_2$ such that $a \neq 0$. Let $\mathbf{b} = [1 \quad -1]^T$. Then

$$\mathbf{a} \odot \mathbf{b} = circ\,(\mathbf{a})\,\mathbf{b} = [0 \quad 0]^T.$$

Similar examples can be constructed for any n. However, as a commutative ring with unity, $\mathbb{K}_n$ does have enough structure to allow the construction of a *module*, if not a vector space.

Definition 3.4. Let $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ denote the space of $n \times n$ matrices with real entries, together with scalars taken from $\mathbb{K}_n$ and with *scalar multiplication* $\circ$ defined as follows:

$$\text{For } \mathbf{a} \in \mathbb{K}_n, \quad X \in \mathbb{R}^{n\times n}, \quad \mathbf{a} \circ X \equiv X\, circ\, (\mathbf{a})$$

where the operation on the right-hand side is the usual matrix–matrix multiplication. When context permits, we will omit $\circ$ and use juxtaposition.

Theorem 3.5. *$\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is a unitary module over $\mathbb{K}_n$.*

Proof. Let $\mathbf{a}, \mathbf{b} \in \mathbb{K}_n$ and $X, Y \in \mathbb{R}^{n\times n}$. Then it suffices to show:

1. $\mathbf{a}(X+Y) = \mathbf{a}X + \mathbf{b}Y$
2. $(\mathbf{a}+\mathbf{b})X = \mathbf{a}X + \mathbf{b}X$
3. $\mathbf{a} \circ (\mathbf{b} \circ X) = (\mathbf{a} \odot \mathbf{b}) \circ X$ and
4. There exists $\mathbf{e} \in \mathbb{K}_n$ such that $\mathbf{e}X = X$ for all $X \in \mathbb{R}^{n\times n}$.

The proofs are as follows:

1.

$$\begin{aligned}\mathbf{a}(X+Y) &= (X+Y)\, circ\, (\mathbf{a})\\ &= X\, circ\, (\mathbf{a}) + Y\, circ\, (\mathbf{b})\\ &= \mathbf{a}X + \mathbf{b}Y\end{aligned}$$

2.

$$\begin{aligned}(\mathbf{a}+\mathbf{b})X &= X\, circ\, (\mathbf{a}+\mathbf{b})\\ &= X(circ\, (\mathbf{a}) + circ\, (\mathbf{b}))\\ &= X\, circ\, (\mathbf{a}) + X\, circ\, (\mathbf{b})\\ &= \mathbf{a}X + \mathbf{b}X\end{aligned}$$

3. Let $F = circ\, (\mathbf{b})\, circ\, (\mathbf{a})$ and let $G = circ\, (\mathbf{a} \odot \mathbf{b}) = circ\, (circ\, (\mathbf{a})\, \mathbf{b})$. In order to show $F = G$, since these are both circulant matrices, it suffices to show that $F_{(:,1)} = G_{(:,1)}$. By the definition of *circ*,

$$F_{(:,1)} = circ\, (\mathbf{b})\, \mathbf{a} \quad \text{and} \quad G_{(:,1)} = circ\, (\mathbf{a})\, \mathbf{b}.$$

Denote the columns of $circ\,(\mathbf{b})$ by $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$, and let $\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}^T$. Then

$$\begin{aligned}F_{(:,1)} &= \left[\ \mathbf{b}_1 \ \middle|\ \mathbf{b}_2 \ \middle|\ \cdots \ \middle|\ \mathbf{b}_n\ \right] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\\ &= a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n\\ &= \begin{bmatrix} a_1b_1 + a_2b_n + \cdots + a_nb_2 \\ a_1b_2 + a_2b_1 + \cdots + a_nb_n \\ \vdots \\ a_1b_n + a_2b_{n-1} + \cdots + a_nb_1 \end{bmatrix}\end{aligned}$$

$$= \begin{bmatrix} a_1b_1 + a_nb_2 + \cdots + a_2b_n \\ a_2b_1 + a_1b_2 + \cdots + a_3b_n \\ \vdots \\ a_nb_1 + a_{n-1}b_2 + \cdots + a_1b_n \end{bmatrix}$$
$$= circ\,(\mathbf{a})\,\mathbf{b} = G_{(:,1)}.$$

Thus, $F = G$. Specifically, that means,

$$circ\,(\mathbf{b})\ circ\,(\mathbf{a}) = circ\,(circ\,(\mathbf{a}))\,\mathbf{b}$$

and so, for any $X \in \mathbb{R}^{n\times n}$,

$$\begin{aligned} &X\,circ\,(\mathbf{b})\ circ\,(\mathbf{a}) = X\,circ\,(circ\,(\mathbf{a})\,\mathbf{b}) \\ &\mathbf{a}\ \circ [X\,circ\,(\mathbf{b})] = [circ\,(\mathbf{a})\,\mathbf{b}]\ \circ X \\ &\mathbf{a}\ \circ (\mathbf{b}\ \circ X) = (\mathbf{a} \odot \mathbf{b})\ \circ X. \end{aligned}$$

4. Let $\mathbf{e} \in \mathbb{K}_n$ be defined such that $\mathbf{e} = (1, 0, 0, \ldots, 0)$. Then $circ\,(\mathbf{e}) = I_n$ and clearly,

$$\mathbf{e}X = X\,circ\,(\mathbf{e}) = X \quad \text{for all } X \in \mathbb{R}^{n\times n}. \qquad \square$$

Theorem 3.6. $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ *has a linearly independent generating set*.

Proof. The proof is by construction. Let B_i denote the $n \times n$ matrix consisting of all zeros except for a 1 in the $(i, 1)$ position. We will show that

$$\mathfrak{B} \equiv \{B_i\,|\,i = 1, 2, \ldots, n\}$$

is a linearly independent generating set for $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$.

1. Let $X \in \mathbb{R}^{n\times n}$. Then define

$$\mathbf{a}_i = \begin{bmatrix} X_{(i,1)} & X_{(i,n)} & X_{(i,n-1)} & X_{(i,n-2)} & \ldots & X_{(i,2)} \end{bmatrix}^T.$$

It is straightforward to verify that

$$\mathbf{a}_1\ \circ B_1 + \mathbf{a}_2\ \circ B_2 + \cdots + \mathbf{a}_n\ \circ B_n = X$$

and so $\mathfrak{B}$ is a generating set.

2. Suppose for some $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$,

$$\mathbf{a}_1\ \circ B_1 + \mathbf{a}_2\ \circ B_2 + \cdots + \mathbf{a}_n\ \circ B_n = \mathbf{0}_n$$

Then, note that, for $1 \leqslant i \leqslant n$, by the structure of B_i, $\mathbf{a}_i \circ B_i$ is a matrix with zero entries in all locations except in the ith row. In that row its entries are the same as the ith row of $circ\,(\mathbf{a})$. This implies $\mathbf{a}_i = \begin{bmatrix} 0 & 0 & \ldots & 0 \end{bmatrix}^T$ for $i = 1, \ldots, n$ and so $\mathfrak{B}$ is linearly independent. $\square$

Now, since $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ has a linearly independent generating set, it is a free module over $\mathbb{K}_n$ (see, for example, [5, Thm IV 2.1]) and, since $\mathbb{K}_n$ is a commutative ring with unity, it has the invariant dimension property ([5, Thm IV 2.12]). This means that any two bases of a free module over $\mathbb{K}_n$ have the same cardinality. Thus, even though $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is not a vector space (since $\mathbb{K}_n$ is not a field), it is, in some sense, the next best thing. We still have the important notions of *linear independence*, *basis* and *subspace*, though the appropriate terminology for modules is *module homomorphism*, *generating set*, and *submodule*, respectively. Most basic properties of vector spaces still hold, though the following example demonstrates one which does not.

Example 3.7. As was shown in Example 3.3, $\mathbb{K}_2$ has zero divisors. This problem persists in $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$. In particular, let $\mathbf{a} = [a \quad a]^T$ with $a \neq 0$ and consider $X \in \mathbb{R}^{2\times 2}$ such that

$$X = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Then

$$\mathbf{a} \circ X = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a & a \\ a & a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

even though neither $\mathbf{a} = 0$ nor $X = \mathbf{0}$.

4. Linearity

Despite the fact that $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is not quite a vector space, it is the proper context for tensor–matrix multiplication as evidenced by the next two theorems.

Theorem 4.1. *Tensor–matrix multiplication $*$, as defined in Definition 2.5, is a linear operator on $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$.*

Proof. Let $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$, $X, Y \in \mathbb{R}^{n\times n}$, and $\mathbf{m} \in \mathbb{K}_n$.

1.
$$\begin{aligned} \mathcal{A} * (X+Y) &= \texttt{fold}\,(circ\,(\mathcal{A})\,vec(X+Y)) \\ &= \texttt{fold}\,(circ\,(\mathcal{A})\,[vec(X)+vec(Y)]) \\ &= \texttt{fold}\,(circ\,(\mathcal{A})\,vec(X)+circ\,(\mathcal{A})\,vec(Y)) \\ &= \texttt{fold}\,(circ\,(\mathcal{A})\,vec(X))+\texttt{fold}\,(circ\,(\mathcal{A})\,vec(Y)) \\ &= \mathcal{A}*X+\mathcal{A}*Y. \end{aligned}$$

2. It remains to show $\mathcal{A} * (\mathbf{m} \circ X) = \mathbf{m} \circ (\mathcal{A} * X)$. Since each side of this statement is an $n \times n$ matrix, we will show equality by considering the ith columns.
First, let $F = \mathbf{m} \circ X$. Then

$$\mathcal{A} * (\mathbf{m} \circ X) = \mathcal{A} * F$$

and, as in Definition 2.5,

$$(\mathcal{A} * F)_{:i} = \sum_{j=1}^{n} A_{row(i,j)} f_j.$$

Since $F = \mathbf{m} \circ X = X\,circ\,(\mathbf{m})$,

$$\begin{aligned} f_j &= X\,(circ\,(\mathbf{m}))_{:j} \\ &= \sum_{k=1}^{n} x_k m_{col(j,k)}. \end{aligned}$$

Then the ith column of $\mathcal{A} * (\mathbf{m} \circ X)$ is

$$(\mathcal{A} * (\mathbf{m} \circ X))_{:i} = \sum_{j=1}^{n} A_{row(i,j)} \left[\sum_{k=1}^{n} x_k m_{col(j,k)}\right] \tag{4a}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{n} A_{row(i,j)} x_k m_{col(j,k)} \tag{4b}$$

$$= \sum_{k=1}^{n}\left[\sum_{j=1}^{n} m_{col(j,k)} A_{row(i,j)}\right] x_k. \tag{4c}$$

Next, consider $\mathbf{m} \circ (\mathcal{A} * X)$. Since $\mathbf{m} \circ (\mathcal{A} * X) = (\mathcal{A} * X)\, circ\,(\mathbf{m})$,

$$(\mathbf{m} \circ (\mathcal{A} * X))_{:i} = \sum_{j=1}^{n} (\mathcal{A} * X)_{:j}\ m_{col(i,j)}.$$

But

$$(\mathcal{A} * X)_{:j} = \sum_{k=1}^{n} A_{row(j,k)} x_k,$$

and so the ith column of $\mathbf{m} \circ (\mathcal{A} * X)$ is

$$(\mathbf{m} \circ (\mathcal{A} * X))_{:i} = \sum_{j=1}^{n} \left[\sum_{k=1}^{n} A_{row(j,k)} x_k \right] m_{col(i,j)} \tag{5a}$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} m_{col(i,j)} A_{row(j,k)} x_k \tag{5b}$$

$$= \sum_{k=1}^{n} \left[\sum_{j=1}^{n} m_{col(i,j)} A_{row(j,k)} \right] x_k. \tag{5c}$$

Finally, it remains to be shown that the inner sums from Eqs. (4c) and (5c) are equal. This follows from the structure of the circulant matrices involved. First, note that, for any $j, k \in \mathbb{N}$,

$$col(j,k) = row(k,j) \quad \text{and} \quad row(j,k) = col(k,j).$$

So, the sums over j from Eqs. (4c) and (5c) can be written as

$$\sum_{j=1}^{n} m_{row(k,j)} A_{row(i,j)} \quad \text{and} \quad \sum_{j=1}^{n} m_{col(i,j)} A_{col(k,j)},$$

respectively. Then, for fixed i and k, the equality of those sums follows from the structure of $circ\,(\mathbf{m})$ and $circ\,(\mathcal{A})$.

Theorem 4.2. *Let $T : \mathbb{R}^{n \times n}_{\mathbb{K}_n} \to \mathbb{R}^{n \times n}_{\mathbb{K}_n}$ be a linear transformation. Then there exists a tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times n}$ such that*

$$T(X) = \mathcal{A} * X \quad \text{for any matrix } X \in \mathbb{R}^{n \times n}.$$

Proof. Recall the linearly independent generating set for $\mathbb{R}^{n \times n}_{\mathbb{K}_n}$ from Theorem 3.6:

$$\mathfrak{B} \equiv \{B_i \,|\, i = 1, 2, \ldots, n\}$$

where B_i denotes the $n \times n$ matrix consisting of all zeros except for a 1 in the $(i, 1)$ position. Then construct an $n \times n \times n$ tensor $\mathcal{A}$ *lateral-slice-wise* by letting

$$\mathcal{A}_{:i:} = T(B_i) \quad \text{for } i = 1, \ldots, n.$$

Let $X \in \mathbb{R}^{n \times n}$. Then, since $\mathfrak{B}$ is a generating set, there exists $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_n \in \mathbb{K}_n$ such that

$$X = \mathbf{m}_1 B_1 + \mathbf{m}_2 B_2 + \cdots + \mathbf{m}_n B_n.$$

Then

$$\begin{aligned} \mathcal{A} * X &= \mathcal{A} * (\mathbf{m}_1 B_1 + \mathbf{m}_2 B_2 + \cdots + \mathbf{m}_n B_n) \\ &= \mathbf{m}_1 (\mathcal{A} * B_1) + \mathbf{m}_2 (\mathcal{A} * B_2) + \cdots + \mathbf{m}_n (\mathcal{A} * B_n) \\ &= \mathbf{m}_1 \mathcal{A}_{:1:} + \mathbf{m}_2 \mathcal{A}_{:2:} + \cdots + \mathbf{m}_n \mathcal{A}_{:n:} \end{aligned}$$

$$
\begin{aligned}
&= \mathbf{m}_1 T(B_1) + \mathbf{m}_2 T(B_2) + \cdots + \mathbf{m}_n T(B_n) \\
&= T(\mathbf{m}_1 B_1 + \mathbf{m}_2 B_2 + \cdots + \mathbf{m}_n B_n) \\
&= T(X)
\end{aligned}
$$

And so, $*$-multiplication by tensor $\mathcal{A}$ represents the linear transformation T. $\square$

Although the space $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is certainly more complicated than the usual vector space of $n \times n$ matrices, Theorem 4.2 shows that *every* linear transformation on $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ can be represented by tensor multiplication and, consequently, many of our familiar techniques of matrix algebra will generalize to tensor algebra.

5. Tensor diagonalization

The motivation for the construction of the free module $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ came from the desire to understand the action of $n \times n \times n$ tensors upon $n \times n$ matrices in light of the tensor diagonalization presented in Eq. (1). Theorem 4.2 shows that $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is the appropriate space in which to study this diagonalization and the corresponding invariant subspaces. This claim is further supported by the following observation.

Theorem 5.1. *Let $\mathcal{A}$, $\mathcal{X}$ and $\mathcal{L}$ be $n \times n \times n$ tensors such that*

$$\mathcal{A} * \mathcal{X} = \mathcal{X} * \mathcal{L}$$

Then, for each $1 \leqslant i \leqslant n$, there exist $\mathbf{m}_i \in \mathbb{K}_n$ such that

$$\mathcal{A} * (\mathcal{X})_{:i:} = \mathbf{m}_i \circ (\mathcal{X})_{:i:}. \tag{6}$$

Specifically,

$$\mathbf{m}_i = \begin{bmatrix} \mathcal{L}_{ii1} & \mathcal{L}_{iin} & \mathcal{L}_{ii(n-1)} & \dots & \mathcal{L}_{ii2} \end{bmatrix}^T.$$

Adapting our familiar matrix terminology, this theorem could be restated that each lateral slice of $\mathcal{X}$ is an *eigenmatrix* $\mathcal{X}_{:i:}$ of $\mathcal{A}$ with corresponding *eigenvector* $\mathbf{m}_i$. However, since the *eigenvector* is now playing the part of the *scalar* in Eq. (6), it may be desirable to use a name less likely to be confused with our previous understanding of eigenvectors. Towards this end we shall refer to the $\mathbf{m}_i$ from Eq. (6) as *eigentuples*.

The proof consists of considering the ith lateral slices of each side of Eq. (1), expanding the right-hand side and then using the simplification afforded by the f-diagonal structure of $\mathcal{L}$.

Proof. For any $1 \leqslant i \leqslant n$, consider the ith lateral slices of each side of Eq. (1). Then, by Lemma 2.7,

$$\mathcal{A} * (\mathcal{X})_{:i:} = \mathcal{X} * (\mathcal{L})_{:i:}.$$

Expanding the right-hand side we see

$$\mathcal{X} * (\mathcal{L})_{:i:} = \texttt{fold}\left(\begin{bmatrix} \mathcal{X}_{::1} & \mathcal{X}_{::n} & \cdots & \mathcal{X}_{::2} \\ \mathcal{X}_{::2} & \mathcal{X}_{::1} & \cdots & \mathcal{X}_{::3} \\ \vdots & \vdots & & \vdots \\ \mathcal{X}_{::n} & \mathcal{X}_{::(n-1)} & \cdots & \mathcal{X}_{::1} \end{bmatrix} \begin{bmatrix} \mathcal{L}_{:i1} \\ \mathcal{L}_{:i2} \\ \vdots \\ \mathcal{L}_{:in} \end{bmatrix} \right).$$

Since $(\mathcal{L})_{:i:}$ is a matrix, as in the remarks following Definition 2.5, the jth column of the result is

$$[\mathcal{X} * (\mathcal{L})_{:i:}]_{:j} = \sum_{k=1}^{n} \mathcal{X}_{::row(j,k)} \mathcal{L}_{:ik}$$

Using the f-diagonal structure of $\mathcal{L}$, each vector $\mathcal{L}_{:ik}$ has only one non-zero entry $\mathcal{L}_{iik}$ in row i. Consequently,

$$[\mathcal{X} * (\mathcal{L}_{:i:})]_{:j} = \sum_{k=1}^{n} \mathcal{L}_{iik} \mathcal{X}_{:i\,row(j,k)}$$
$$= \sum_{k=1}^{n} \mathcal{L}_{ii\,row(j,k)} \mathcal{X}_{:ik}.$$

Finally, we have

$$\mathcal{X} * (\mathcal{L})_{:i:} = \left[\ \mathcal{X}_{:i1} \mid \mathcal{X}_{:i2} \mid \cdots \mid \mathcal{X}_{:in}\ \right] \begin{bmatrix} \mathcal{L}_{ii1} & \mathcal{L}_{ii2} & \cdots & \mathcal{L}_{iin} \\ \mathcal{L}_{iin} & \mathcal{L}_{ii1} & \cdots & \mathcal{L}_{ii(n-1)} \\ \vdots & \vdots & & \vdots \\ \mathcal{L}_{ii2} & \mathcal{L}_{ii3} & \cdots & \mathcal{L}_{ii1} \end{bmatrix},$$

and so

$$\mathcal{A} * (\mathcal{X})_{:i:} = \mathbf{m}_i \circ (\mathcal{X})_{:i:}$$

where $\mathbf{m}_i = \left[\mathcal{L}_{ii1} \quad \mathcal{L}_{iin} \quad \mathcal{L}_{ii(n-1)} \quad \ldots \quad \mathcal{L}_{ii2}\right]^T$. $\square$

While our goal has been to provide a framework for understanding the action of $n \times n \times n$ tensors upon $n \times n$ matrices, it remains to consider the existence of a factorization as presented in Eq. (1). To begin, let $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$ and recall that *circ* $(\mathtt{unfold}(\mathcal{A}))$ is a block circulant matrix. Then, following the approach of Kilmer et al. [4], we note that such a matrix can be block diagonalized by means of the Discrete Fourier Transform. That is,

$$(F_n \otimes I_n)\, circ\, (\mathtt{unfold}(\mathcal{A}))\, (F_n^* \otimes I_n) = \begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_n \end{bmatrix} \tag{7}$$

where F_n is the $n \times n$ DFT matrix, I_n is the $n \times n$ identity, F_n^* denotes the conjugate transpose of F_n and $\otimes$ indicates the Kronecker product. The crux of the existence of Eq. (1) lies in the diagonalizability of the D_i's. If each D_i is diagonalizable, then tensors $\mathcal{X}$ and $\mathcal{L}$ can be constructed using the same process as used in [4] to produce the tensor SVD.

6. Conclusions and further research

In an effort to further an understanding of "the objects that are operated on by tensors" [1,3] and to provide a useful framework in which to view the action, we have constructed the free module $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$. The advantage of this frame of reference is that all linear transformations on $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ can be represented by tensor–tensor multiplication. The definition of $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ is the first step in extending many of our familiar matrix linear algebra tools to a tensor linear algebra setting. For example, Theorem 4.2 allows for an interpretation of Eq. (1) as a statement encapsulating invariant submodule information in the form of eigenmatrices and eigentuples. Indeed, the "eigenvalues" of a tensor are vectors rather than scalars! [1].

While providing answers to some questions, this work in fact opens many more. One which merits further attention is whether the ring $\mathbb{K}_n$ is, modulo some ideal, isomorphic to a field. We are currently pursuing this idea, but preliminary results indicate that such an isomorphism is not possible. Also under consideration is the relationship of $\mathbb{K}_n$ to the n-complex numbers described by Olariu in [6].

At first glance the eigenmatrices and eigentuples presented here are inherently different from the tensor eigenvalues described by Lim [7] and Qi [8], because the operation of the tensors is defined in a new way. However, it may be that, with further work, a relationship between the two different approaches may be discovered.

Additional areas of interest include extending our understanding of $\mathbb{R}^{n\times n}_{\mathbb{K}_n}$ and investigating further analogs of both matrix decompositions. For instance, is there a natural way to define an inner product on this space? With the newly defined tensor–tensor and tensor–matrix multiplications and the resulting factorizations, does there exist a tensor analog to the Jordan Canonical Form? Is an LU factorization possible and what information could it provide? Many of these are among the questions originally generated at the "Workshop on Tensor Decompositions" [1].

Finally, it remains to be seen what new insights and challenges will arise from the interplay of $\mathbb{R}_{\mathbb{K}_n}^{n\times n}$ and the tools and techniques of numerical linear algebra. Reliable, efficient computation of these new tensor factorizations will be of interest, as will extensions of algorithms such as Gram–Schmidt orthogonalization, the QR algorithm and Krylov subspace methods.

References

[1] C.D.M. Martin, Tensor decompositions workshop discussion notes, American Institute of Mathematics, http://www.aimath.org/WWN/tensordecomp/tensordecomp.pdf (accessed at August, 2008), July 2004.
[2] American Institute of Mathematics, http://www.aimath.org/WWN/tensordecomp/ (accessed at August, 2008), July 2004.
[3] M.E. Kilmer, C.D. Moravitz Martin, Decomposing a tensor, SIAM News 39 (9) (2004).
[4] M.E. Kilmer, C.D. Martin, L. Perrone, A third-order Generalization of the Matrix SVD as a Product of Third-order Tensors, Tech. Rep. TR-2008-4, Tufts University, Department of Computer Science, October 2008.
[5] T.W. Hungerford, Algebra, Graduate Texts in Mathematics, vol. 73, Springer-Verlag, New York, 1974.
[6] S. Olariu, Complex Numbers in n Dimensions, Elsevier, Amsterdam, 2002.
[7] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach, in: Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processes, vol. 1, 2005, pp. 120–132.
[8] L. Qi, Eigenvalues of a real supersymmetric tensor, J. Symbolic Comput. 40 (6) (2005) 1302–1324. doi:10.1016/j.jsc.2005.05.007.